Deep-learning technique predicts clinical treatment outcomes

February 25 2022, by Lauren Hinkel

Credit: Unsplash/CC0 Public Domain

When it comes to treatment strategies for critically ill patients, clinicians want to be able to consider all their options and timing of administration, and make the optimal decision for their patients. While clinician experience and study has helped them to be successful in this effort, not all patients are the same, and treatment decisions at this crucial time could mean the difference between patient improvement and quick deterioration. Therefore, it would be helpful for doctors to be able to take a patient's previous known health status and received treatments and use that to predict that patient's health outcome under different treatment scenarios, in order to pick the best path.

Now, a deep-learning technique, called G-Net, from researchers at MIT and IBM provides a window into causal counterfactual prediction, affording physicians the opportunity to explore how a patient might fare under different treatment plans. The foundation of G-Net is the g-computation algorithm, a causal inference method that estimates the effect of dynamic exposures in the presence of measured confounding variables—ones that may influence both treatments and outcomes. Unlike previous implementations of the g-computation framework, which have used linear modeling approaches, G-Net uses recurrent neural networks (RNN), which have node connections that allow them to better model temporal sequences with complex and nonlinear dynamics, like those found in the physiological and clinical time series data. In this way, physicians can develop alternative plans based on patient history and test them before making a decision.

"Our ultimate goal is to develop a machine learning technique that would allow doctors to explore various "What if' scenarios and treatment options," says Li-wei Lehman, MIT research scientist in the MIT Institute for Medical Engineering and Science and an MIT-IBM Watson AI Lab project lead. "A lot of work has been done in terms of deep learning for counterfactual prediction but [it's] been focusing on a point exposure setting," or a static, time-varying treatment strategy, which doesn't allow for adjustment of treatments as patient history changes. However, her team's new prediction approach provides for treatment plan flexibility and chances for treatment alteration over time as patient covariate history and past treatments change. "G-Net is the first deep-learning approach based on g-computation that can predict both the population-level and individual-level treatment effects under dynamic and time varying treatment strategies."

The research, which was recently published in *Proceedings of Machine Learning Research*, was co-authored by Rui Li MEng '20, Stephanie Hu MEng '21, former MIT postdoc Mingyu Lu MD, graduate student Yuria

Utsumi, IBM research staff member Prithwish Chakraborty, IBM Research director of Hybrid Cloud Services Daby Sow, IBM data scientist Piyush Madan, IBM research scientist Mohamed Ghalwash, and IBM research scientist Zach Shahn.

Tracking disease progression

To build, validate, and test G-Net's predictive abilities, the researchers considered the circulatory system in septic patients in the ICU. During critical care, doctors need to make trade-offs and judgment calls, such as ensuring the organs are receiving adequate blood supply without overworking the heart. For this, they could give intravenous fluids to patients to increase blood pressure; however, too much can cause edema. Alternatively, physicians can administer vasopressors, which act to contract blood vessels and raise blood pressure.

In order to mimic this and demonstrate G-Net's proof-of-concept, the team used CVSim, a mechanistic model of a human cardiovascular system that's governed by 28 input variables characterizing the system's current state, such as arterial pressure, central venous pressure, total blood volume, and total peripheral resistance, and modified it to simulate various disease processes (e.g., sepsis or blood loss) and effects of interventions (e.g., fluids and vasopressors). The researchers used CVSim to generate observational patient data for training and for "ground truth" comparison against counterfactual prediction. In their G-Net architecture, the researchers ran two RNNs to handle and predict variables that are continuous, meaning they can take on a range of values, like blood pressure, and categorical variables, which have discrete values, like the presence or absence of pulmonary edema. The researchers simulated the health trajectories of thousands of "patients" exhibiting symptoms under one treatment regime, let's say A, for 66 timesteps, and used them to train and validate their model.

Testing G-Net's prediction capability, the team generated two counterfactual datasets. Each contained roughly 1,000 known patient health trajectories, which were created from CVSim using the same "patient" condition as the starting point under treatment A. Then at timestep 33, treatment changed to plan B or C, depending on the dataset. The team then performed 100 prediction trajectories for each of these 1,000 patients, whose treatment and medical history was known up until timestep 33 when a new treatment was administered. In these cases, the prediction agreed well with the "ground-truth" observations for individual patients and averaged population-level trajectories.

A cut above the rest

Since the g-computation framework is flexible, the researchers wanted to examine G-Net's prediction using different nonlinear models—in this case, long short-term memory (LSTM) models, which are a type of RNN that can learn from previous data patterns or sequences—against the more classical linear models and a multilayer perception model (MLP), a type of neural network that can make predictions using a nonlinear approach. Following a similar setup as before, the team found that the error between the known and predicted cases was smallest in the LSTM models compared to the others. Since G-Net is able to model the temporal patterns of the patient's ICU history and past treatment, whereas a linear model and MLP cannot, it was better able to predict the patient's outcome.

The team also compared G-Net's prediction in a static, time-varying treatment setting against two state-of-the-art deep-learning based counterfactual prediction approaches, a recurrent marginal structural network (rMSN) and a counterfactual recurrent neural network (CRN), as well as a linear model and an MLP. For this, they investigated a model for tumor growth under no treatment, radiation, chemotherapy, and both radiation and chemotherapy scenarios. "Imagine a scenario where there's

a patient with cancer, and an example of a static regime would be if you only give a fixed dosage of chemotherapy, radiation, or any kind of drug, and wait until the end of your trajectory," comments Lu. For these investigations, the researchers generated simulated observational data using tumor volume as the primary influence dictating treatment plans and demonstrated that G-Net outperformed the other models. One potential reason could be because g-computation is known to be more statistically efficient than rMSN and CRN, when models are correctly specified.

While G-Net has done well with simulated data, more needs to be done before it can be applied to real patients. Since neural networks can be thought of as "black boxes" for prediction results, the researchers are beginning to investigate the uncertainty in the model to help ensure safety. In contrast to these approaches that recommend an "optimal" treatment plan without any clinician involvement, "as a decision support tool, I believe that G-Net would be more interpretable, since the clinicians would input treatment strategies themselves," says Lehman, and "G-Net will allow them to be able to explore different hypotheses." Further, the team has moved on to using real data from ICU patients with sepsis, bringing it one step closer to implementation in hospitals.

"I think it is pretty important and exciting for real-world applications," says Hu. "It'd be helpful to have some way to predict whether or not a treatment might work or what the effects might be—a quicker iteration process for developing these hypotheses for what to try, before actually trying to implement them in in a years-long, potentially very involved and very invasive type of clinical trial."

More information: Rui Li et al, G-Net: a Recurrent Network Approach to G-Computation for Counterfactual Prediction Under a Dynamic Treatment Regime, *Proceedings of Machine Learning Research* (2021). proceedings.mlr.press/v158/li21a.html

This story is republished courtesy of MIT News (web.mit.edu/newsoffice/), a popular site that covers news about MIT research, innovation and teaching.

Provided by Massachusetts Institute of Technology